Figure 1.

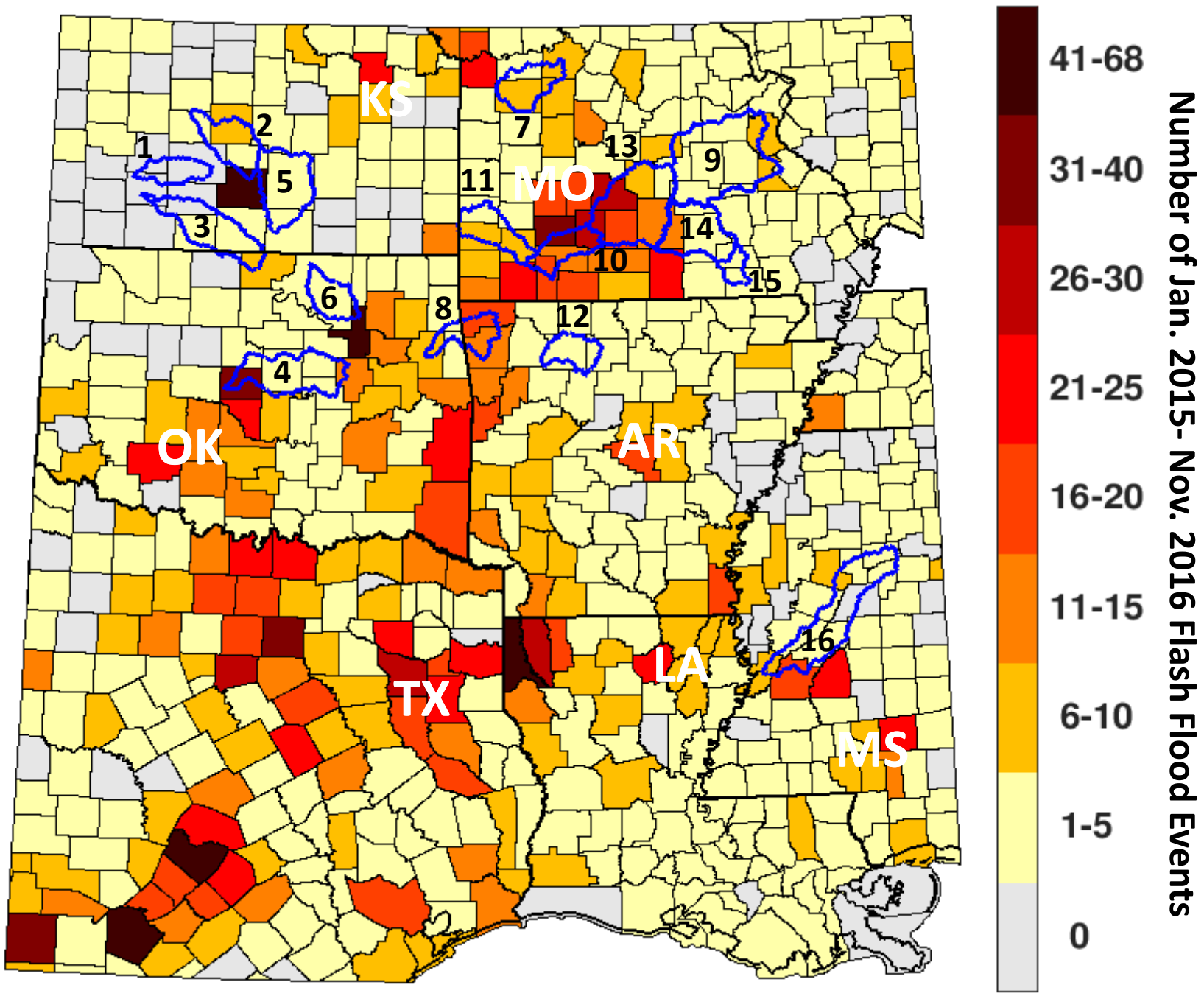

KS
MO
AR
OK
TX
LA
MS
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
41-68
31-40
26-30
21-25
16-20
11-15
6-10
1-5
0
Number of Jan. 2015- Nov. 2016 Flash Flood Events

Figure 2.

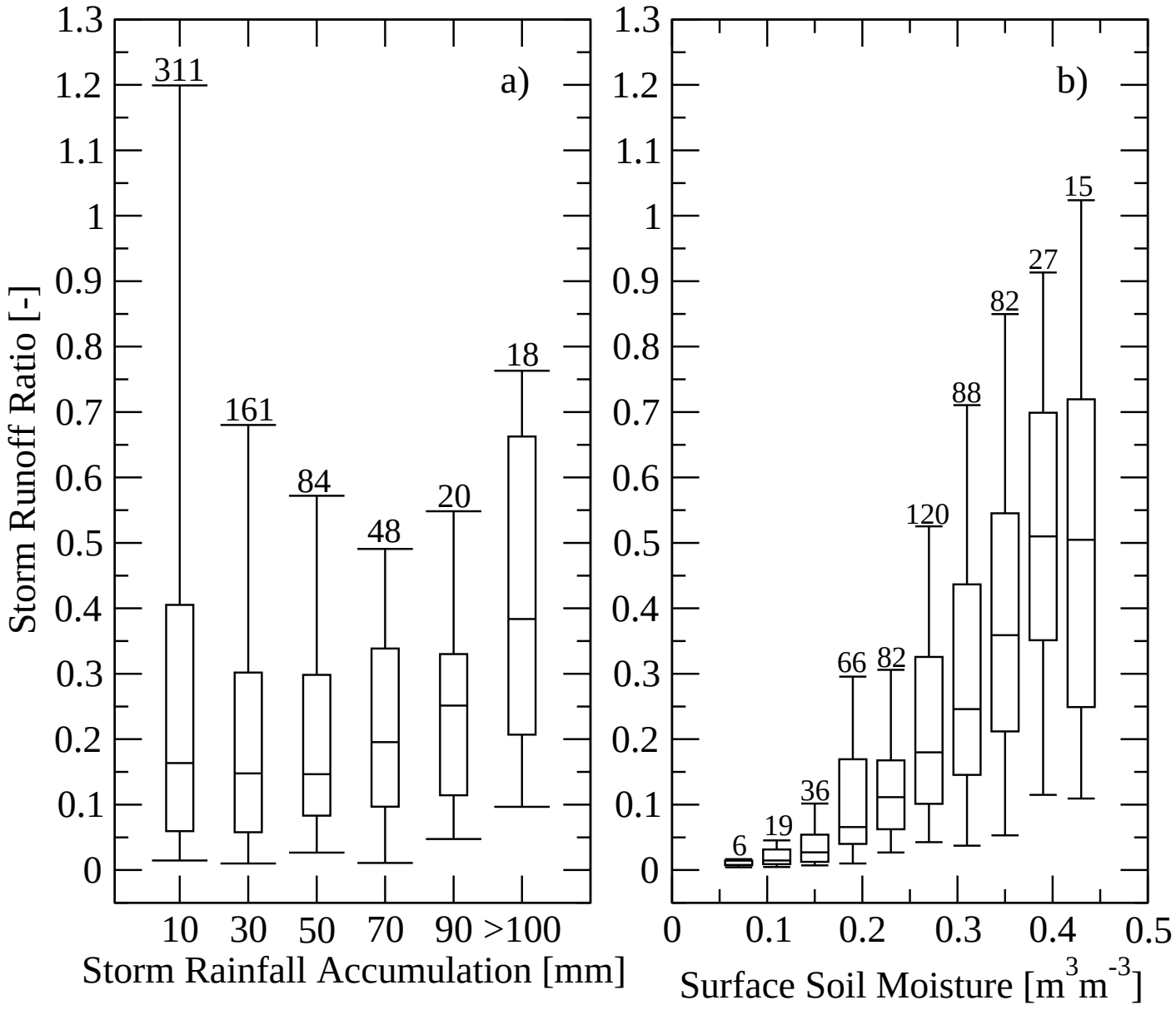
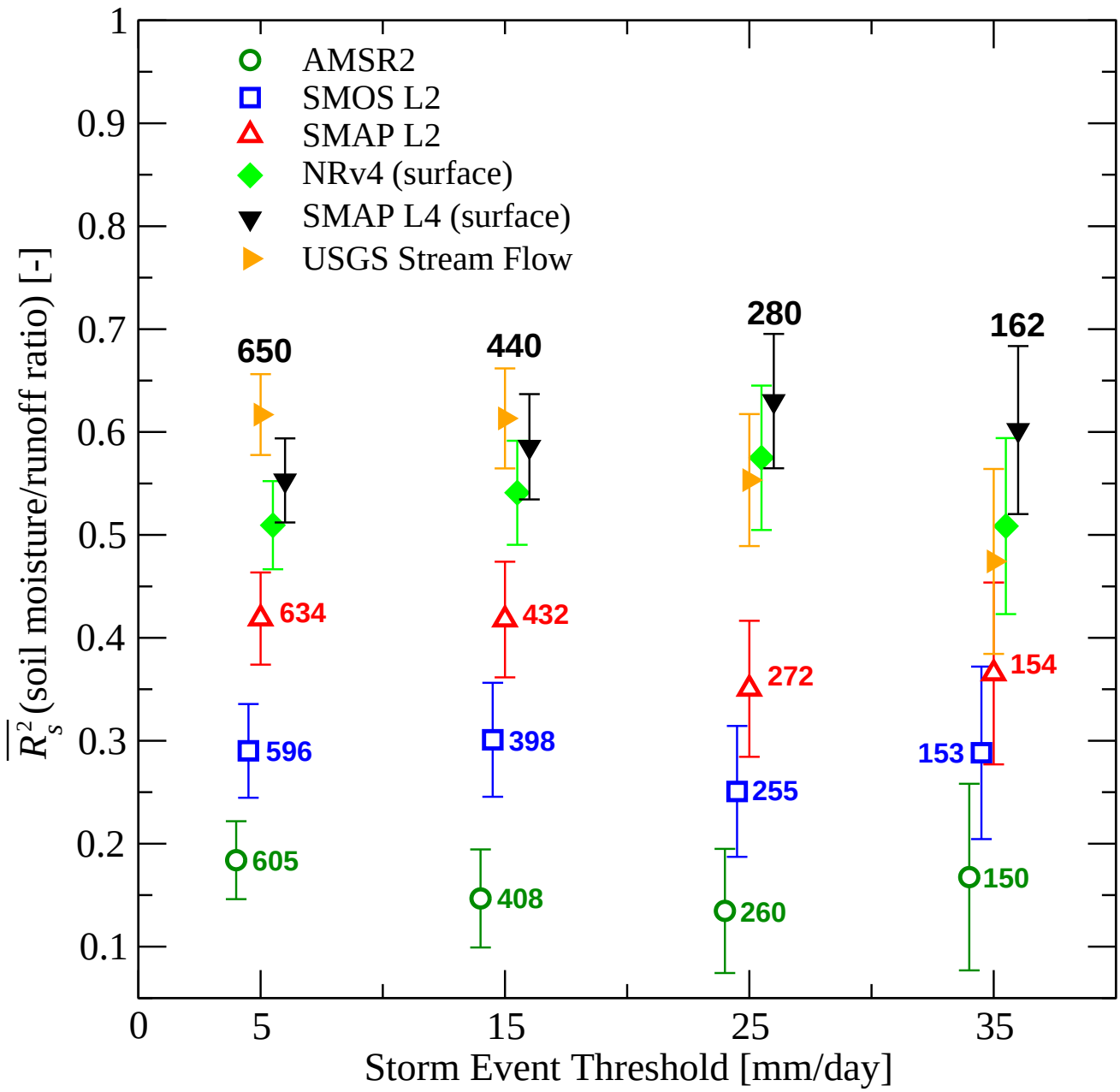

a)
b)
Storm Runoff Ratio [-]
Storm Rainfall Accumulation [mm]
Surface Soil Moisture [m3m-3]
311
161
84
48
20
18
6
19
36
66
82
120
88
82
27
15
10
30
50
70
90
>100
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3

Figure 3.


AMSR2
SMOS L2
SMAP L2
NRv4 (surface)
SMAP L4 (surface)
USGS Stream Flow
650
440
280
162
634
432
272
154
596
398
255
153
605
408
260
150
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
5
15
25
35
$\overline{R_s^2}$ (soil moisture/runoff ratio) [-]
Storm Event Threshold [mm/day]

L-band microwave remote sensing and land data assimilation improve the representation of pre-storm soil moisture conditions for hydrologic forecasting

W.T. Crow[1], F. Chen[1,3], R.H. Reichle[2], and Q. Liu[2,3]

[1]USDA Hydrology and Remote Sensing Laboratory, Beltsville, MD

[2]NASA GSFC Global Modeling and Assimilation Office, Greenbelt, MD

[3]Science Systems and Applications Inc., Greenbelt, MD

Abstract

Recent advances in remote sensing and land data assimilation purport to improve the quality of antecedent soil moisture information available for operational hydrologic forecasting. We objectively validate this claim by calculating the strength of the relationship between storm-scale runoff ratio (i.e., total stream flow divided by total rainfall accumulation in depth units) and pre-storm surface soil moisture estimates from a range of surface soil moisture data products. Results demonstrate that both satellite-based, L-band microwave radiometry and the application of land data assimilation techniques have significantly improved the utility of surface soil moisture data sets for forecasting stream flow response to future rainfall events.

1. Introduction

Anticipating the capacity of the land surface to infiltrate future rainfall is an important source of predictability in short-term operational stream flow forecasts [Silvestro et al., 2014; Massari et al. 2014]. Dynamic changes in this capacity are due primarily to variations in soil moisture content, which determine the infiltration capacity of the soil column [Western and Grayson,

1998]. As a result, there has been considerable interest in using remotely-sensed surface soil moisture retrievals for improved monitoring of pre-storm soil moisture conditions within hydrologic basins [Massari et al., 2015a]. However, these retrievals suffer from a number of well-known weaknesses including: 1) coarse spatial resolution (typically > 30 km), 2) shallow vertical support within the soil column (typically 1-5 cm), and 3) reduced accuracy under dense vegetation.

Therefore, robust evaluation techniques are needed to objectively measure the benefits of new soil moisture products for hydrologic forecasting. One common approach has been to compare hydrologic model performance before and after the assimilation of a remotely-sensed soil moisture product. However, a review of these approaches reveals a wide disparity in conclusions regarding the value of soil moisture assimilation for forecasting stream flow [Crow and Ryu, 2008; Massari et al., 2015b; Lievens et al., 2015]. This lack of consistency arises, at least in part, from significant sensitivity to the structure and calibration of the particular hydrologic model applied in the assimilation system [Chen et al., 2009; Zhuo and Han, 2016; Massari et al., 2015a]. Therefore, evaluation results are non-robust in that they are affected by the accuracy of the assumed parametric relationship connecting precipitation, runoff and soil moisture imbedded within these models. In order to remove this sensitivity, and provide a more robust basis for cross-comparing a wide range of soil moisture products, Crow et al. [2005] developed a simplified evaluation approach based on temporally sampling the Spearman rank correlation between pre-storm soil moisture and (subsequent) storm-scale runoff ratios – defined as the ratio of total storm-scale stream flow to total storm-scale rainfall accumulation (both in dimensions of length) over a ~1 week period following a triggering precipitation event.

There has been considerable recent progress in the development of operational soil moisture products. These advances include the 2009 launch of the European Space Agency Soil Moisture and Ocean Salinity (SMOS) mission [Kerr et al., 2010] and the 2015 launch of the National Aeronautics and Space Administration Soil Moisture Active Passive (SMAP) mission [Entekhabi et al., 2010], both dedicated to measuring global surface soil moisture using L-band microwave radiometry, as well as the development of operational, value-added soil moisture data products based on the assimilation of L-band observations into a land surface model, such as the SMAP Level 4 Surface and Root-zone Soil Moisture (SMAP_L4) product [Reichle et al., 2016]. Our goal here is to update Crow et al. [2005] to consider these new soil moisture products and provide an objective description of their relative value for hydrologic forecasting.

2. Study basins and data

This study focuses on 16 medium-scale (2,000-10,000 km^2) hydrologic basins located within the south-central United States (Figure 1). This particular region has experienced an unusually large number of flash flooding events during the past two years (Figure 1) and is therefore a natural choice for an analysis aimed at hydrologic predictability. In addition, land cover conditions in the region are generally amenable to the remote sensing of soil moisture (i.e., there is infrequent snow cover, generally modest topographic relief, and relatively isolated forest coverage). The selection of specific basins within this region was based on a screening analysis performed by the Model Parameterization Experiment [Duan et al., 2006] which identified suitable basins with adequate rain gauge density and lacking significant amounts of anthropogenic impoundment or diversion of stream flow.

Individual basin characteristics are summarized in Table 1. Moving from west to east, these basins exhibit progressively higher mean annual rainfall and runoff ratios (Table 1). Western basins are generally characterized by rangeland, grassland and winter wheat land cover types with relatively low biomass. More easterly basins contain larger amounts of upland forest cover and summer agriculture in low-lying areas.

For each basin, daily rainfall accumulations are derived from the spatial and temporal aggregation of gauge-corrected, 4-km Stage IV precipitation [Lin, 2011] data (to a daily time scale and a basin-average spatial scale) and daily stream flow values based on United States Geological Survey (USGS) stream gauge measurements located at each basin outlet [USGS, 2016]. Rainfall accumulation and stream flow daily totals are computed for 0 to 24 LST (UTC-6 hours). Antecedent soil moisture estimates are obtained from each of the sources described below.

2.1 AMSR2

AMSR2 soil moisture retrievals were based on the application of the Land Parameter Retrieval Model (LPRM) to the ~35-km resolution X-band channel of the Japanese Space Agency Advanced Microwave Scanning Radiometer-2 (AMSR2) satellite sensor to produce a 0.25° resolution product [Vrije Universiteit Amsterdam and NASA GSFC, 2014; Parinussa et al., 2015]. Owing to known problems with LPRM retrievals obtained at the 1:30 PM AMSR2 ascending overpass [Lei et al., 2015], only retrievals from the 1:30 AM descending overpass were utilized. In addition, retrievals with uncertainties greater than 0.40 m^3m^{-3} were masked. These masked retrievals comprise approximately 11% of all AMSR2 retrievals in the study region. The AMSR2 sensor also measures in a (lower frequency) C-band channel which is

92 suitable for retrieving soil moisture; however, this channel is known to be contaminated by radio
93 frequency interference over the United States.

94 2.2 SMOS L2

95 The SMOS mission [Kerr et al., 2010] measures L-band (1.400-1.427 GHz) microwave
96 brightness temperature at ~45-km spatial resolution with equatorial ascending/descending
97 overpasses at approximately 6 am/pm local solar time and a 3-day revisit period at the equator. It
98 began scientific data collection in January 2010. The SMOS Level 2 (L2) soil moisture product
99 utilized here is based on application of SMOS processor version 6.2.0 to retrieve soil moisture on
100 an equal-area ISEA4h9 15-km grid [Kerr et al., 2012]. SMOS_L2 retrievals obtained from both
101 ascending (6 pm) and descending (6 am) orbits were combined into a single time series.
102 Normalized retrieval error was determined by dividing the SMOS data quality index value
103 (provided with each soil moisture value) by the absolute SMOS_L2 soil moisture estimate. All
104 retrievals with normalized error greater than 0.50 [-] were masked from the analysis. These
105 masked retrievals comprise approximately 7% of all SMOS_L2 retrievals in the study region.

106 2.3 SMAP L2

107 Launched in January 2015, SMAP began continuous science data acquisition on March 31, 2015
108 with its L-band (1.41 GHz) radiometer [Entekhabi at al., 2010]. The SMAP Enhanced Level 2
109 (L2) Passive Soil Moisture product is generated by applying the Backus-Gilbert optimal
110 interpolation technique to the original SMAP brightness temperature product and then the SMAP
111 baseline soil moisture retrieval algorithm [O'Neill et al., 2016]. This version of the SMAP_L2
112 product was released in December 2016 and is posted on version 2 of the global cylindrical 9 km
113 Equal-Area Scalable Earth (EASEv2) grid [Brodzik et al., 2012] with a native resolution of ~36
114 km. Retrievals obtained from both ascending (6 pm) and descending (6 am) orbits were

combined into a single time series. Masking was applied to remove retrievals during periods of snow cover or frozen soil.

2.4 SMAP L4 and NRv4

The SMAP_L4 algorithm is an ensemble-based assimilation system built around the NASA Goddard Earth Observing System version 5 (GEOS-5) Catchment land surface model [Koster et al., 2000]. Its primary drivers are SMAP brightness temperature observations and surface meteorological forcing data from the GEOS-5 atmospheric assimilation system, corrected with precipitation observations [Reichle and Liu, 2014]. The algorithm interpolates and extrapolates information from the SMAP observations in time and in space based on the relative uncertainties of the model estimates and the observations. SMAP_L4 data include 3-hourly soil moisture estimates for the "surface" (0-5 cm) and "root zone" (0-100 cm) layers on the 9-km EASEv2 grid [Reichle et al., 2016]. L4 data are available within 2-3 days from the time of observation. The unpublished Nature Run, version 4 (NRv4) data are also generated with the SMAP_L4 system, but configured for a single ensemble member (no perturbations) and without the assimilation of SMAP brightness temperature observations. As a result, NRv4 provides a model-only reference to assess the relative benefit of assimilating SMAP brightness temperature observations.

3. Approach

3.1. Storm event definition

A storm "event" is defined as the 6-day period following a triggering daily precipitation accumulation amount that exceeds a pre-specified threshold. By design, these triggering events always fall on the first day of this event period, and, to avoid the confounding impact of over-lapping storm events, we discard events for which another storm exceeding the threshold occurs

within the event period. Likewise, all events must be preceded by at least one day with a daily precipitation amount below the storm accumulation threshold. All daily soil moisture products are 0 to 24 LST (UTC-6 hours) averages, and pre-storm antecedent soil moisture is defined as the minimum value of daily soil moisture obtained during the two-day period prior to the onset of a storm event. In all cases, at least 25% spatial coverage is required to sample a basin-average soil moisture value.

Daily stream flow observations (in native flow rate dimensions [L^3/T]) are converted into daily depths [L/T] via normalization by basin area. Daily rainfall and stream flow accumulations are then temporally summed for each storm event and a storm-scale runoff-ratio is calculated for each individual event. For a range of daily precipitation storm event thresholds, the Spearman rank coefficient of variation (R^2_s) between antecedent soil moisture and storm scale runoff-ratio is sampled in time for each basin and each soil moisture product. Rank correlation is used because the relationship between antecedent soil moisture and runoff ratio is potentially nonlinear. Owing to the relatively short length of the SMAP data record to date, sampled R^2_s values for individual basins are subject to large random sampling errors, and we currently lack the statistical power to evaluate soil moisture product performance on a basin-by-basin basis. Therefore, we focus only on spatially-averaged values of R^2_s ($\overline{R_s^2}$) acquired across all 16 basins between 31 March 2015 and 31 December 2016.

No attempt was made to isolate storm flow within the overall stream flow time series. Therefore, it is possible for base flow to contribute a non-insignificant fraction of observed storm-scale stream flow response (especially for low storm precipitation thresholds within relatively humid

161 study basins). However, it should be stressed that the presence of base flow does not undermine
162 the interpretation of $\overline{R_s^2}$ as a metric for stream flow forecasting skill. Instead, it simply indicates
163 that a fraction of this forecasting skill is due to the temporal persistence of elevated base flow
164 levels (associated with high soil moisture values) rather than the prediction of land surface
165 response to future precipitation.

166 3.2. Uncertainty description

167 Uncertainty intervals for R^2_s values sampled within individual basins are obtained using a 5000-
168 member boot-strapping approach and then merged to estimate uncertainty intervals for sampled
169 $\overline{R_s^2}$. Based on the averaged spatial correlation sampled between SMAP_L4 basin-averaged,
170 surface soil moisture values (presumed to be the most accurate representation of soil moisture
171 available), and the approach of Bretherton et al. [1999], the 16 basins in Figure 1 contain only
172 7.4 spatially-independent samples. In addition, since $\overline{R_s^2}$ values for each soil moisture product
173 are sampled from a highly-overlapping set of storm events, uncertainty intervals attached to
174 individual products provide a potentially misleading description of the statistical significance of
175 pair-wise differences (since the cross-correlation of sampling errors ensures that the variance of
176 sampling error in pair-wise differences is less than the sum of the sampling error variances for
177 each product individually). Therefore, we further assess the sampling uncertainty in relative
178 comparisons based on the boot-strapping of pair-wise $\overline{R_s^2}$ differences between all soil moisture
179 products - considering only storm events whose antecedent conditions are captured by both
180 members of the soil moisture product pair.

181

182 **4. Results**

Based on sampling across all storm events and all basins, Figure 2 illustrates the range in observed rainfall runoff ratio and its variation as a function of both storm-scale precipitation accumulation (Figure 2a) and pre-storm surface soil moisture (acquired from the SMAP_L4 product; Figure 2b). As expected, a slight increase in runoff ratio is seen with increased storm size in Figure 2a. However, even for relatively large storm events (with > 100 mm of total rainfall accumulation), a wide range of potential storm-scale runoff ratios is observed (Figure 2a). Runoff ratio exhibits a much stronger overall relationship with pre-storm surface soil moisture levels (Figure 2b; provided again by SMAP_L4) - demonstrating the contribution of antecedent soil moisture conditions to hydrologic predictability.

Figure 3 plots $\overline{R_s^2}$ for precipitation storm thresholds ranging from 5 to 35 mm/day and pre-storm soil moisture products. Recall that $\overline{R_s^2}$ is the spatial average of R^2_s sampled individually within each of our 16 study basins. Numerical labels in Figure 3 reflect the number of storm events sampled to acquire plotted values of $\overline{R_s^2}$. The error bars in Figure 3 capture 95% sampling confidence intervals obtained from the boot-strapping approach described above. However, for reasons discussed above, the pair-wise hypothesis tests presented in Table are used as basis of formal conclusions regarding the statistical significance of sampled $\overline{R_s^2}$ differences between products.

Higher values of $\overline{R_s^2}$ in Figure 3 are consistent with an enhanced ability to detect variations in soil moisture which subsequently impact stream flow response to future precipitation. Among the remote sensing products (open symbols in Figure 3), SMAP_L2 demonstrates the best $\overline{R_s^2}$ results, followed by the SMOS_L2 product, and then the X-band AMSR2 retrievals. For the lower

accumulation thresholds (5, 15 and 25 mm/day), both SMOS_L2 and SMAP_L2 differences versus AMSR2 are statistically-significant (two-tailed, 95% confidence; Table 2). Restricting SMAP_L2 and SMOS_L2 retrievals to only the 6 AM or 6 PM overpasses, to better mimic the use of only the 1:30 AM overpass for AMSR2 retrievals, had only a minimal impact on their sampled $\overline{R_s^2}$ results. Therefore, Figure 3 is consistent with the expectation that L-band remote sensing products are more valuable than older products acquired from higher-frequency microwave channels (e.g., X-band). In addition, SMAP_L2 significantly outperforms AMSR2 for the highest event threshold and SMOS_L2 for the lower two thresholds (5 and 15 mm/day). However, the $\overline{R_s^2}$ differences between SMOS_L2 and SMAP_L2 become non-significant for the 15 and 25 mm/day thresholds (Table 2).

Despite its relative superiority versus other remote-sensing products, the SMAP_L2 product still lags behind surface soil moisture estimates obtained from the NRv4 modeling system (Figure 3). Nevertheless, improvement relative to NRv4 is seen when SMAP brightness temperature observations (which form the basis of the SMAP_L2 retrievals) are assimilated into the NRv4 modeling system to produce the SMAP_L4 product. However, the difference between the SMAP_L4 and NRv4 $\overline{R_s^2}$ falls short of 95% confidence (ranging from between 84% and 91% confidence depending on storm event threshold size - see Table 2). Relatively little difference is found in Figure 3 when switching between the use of surface and “root-zone” SMAP_L4 and NRv4 soil moisture products (not shown). However, this may be simply due to the tendency for the Catchment land surface model (used to generate both products) to exhibit relatively strong vertical coupling between its surface and root-zone soil moisture predictions [Kumar et al., 2009].

In addition to soil moisture products, Figure 3 also examines the use of pre-storm USGS daily stream flow data as a predictor of storm-scale runoff ratios. If available, antecedent stream flow measurements are generally assumed to be a valuable predictor of future stream flow magnitudes and commonly assimilated into operational hydrologic models – see e.g., Liu et al. [2016]. However, for precipitation accumulation thresholds of 15 mm/day and above, the SMAP_L4 product outperforms daily USGS stream flow measurements as a leading predictor of storm-scale runoff ratio - at a significance level which reaches 93% confidence for an event threshold of 35 mm/day (see Table 2).

As noted above, several choices underpin our approach for defining discrete rainfall events within a continuous daily rainfall record. In order to determine the impact of these choices, alternative versions of Figure 3 were generated for the cases of: 1) maximum storm lengths of 5 and 7 days (versus the default of 6 days), 2) the use of prior day soil moisture to define antecedent conditions (versus the default of using the minimum soil moisture estimated in the two-day period prior to the storm events), and 3) not masking storm events which are interrupted by the onset of another event (versus the default of masking these events). None of these tested variations changed the qualitative relationships summarized in Figure 3. Another concern is the impact of including snow events on the sampling of $\overline{R_S^2}$ for the NRv4, SMAP_L4 and USGS Stream flow results plotted in Figure 3. However, sub-setting these datasets to include only days with SMAP_L2 retrievals (which have passed a frozen soil and snow cover mask during processing) had no discernible impact on results. Alternative versions of Figure 3 for all cases listed above are shown in the supporting material (Figures S1, S2, S3, S4 and S5).

5. Summary and Conclusions

Within the range of basins studied here, expectations concerning storm-scale rainfall runoff ratios are strongly conditioned by appropriate knowledge of pre-storm soil moisture conditions (Figure 2b). In addition, the development and application of both L-band radiometry and advanced data assimilation systems have significantly improved the quality of soil moisture information available for this purpose (Figure 3, Table 2). In particular, the assimilation of SMAP L-band brightness temperature data in the SMAP_L4 system results in a surface soil moisture product with the highest hydrologic forecasting skill observed to date, and the SMAP_L4 product provides at least as much predictive skill as pre-storm measurements of stream flow (Figure 3). The relative advantages of the SMAP_L4 product grow as the analysis is focused on larger storm events (see the right-hand-side of Figure 3). It should, however, be stressed that this conclusion is based on a single regional study in an area that is relatively well-suited to the remote retrieval of soil moisture. Follow-on work over a wider range of conditions is needed.

In closing, it should be noted that the successful application of satellite-based soil moisture products for hydrologic forecasting also depends on their near-real time availability. SMAP_L2 products are typically available within 24 hours from the time of observation. SMAP_L4 data are available within 2-3 days because of the latency incurred by the use of gauge-based precipitation inputs. However, several options exist for shortening the latency of SMAP_L2 and L4 products, including the short-term forecasting of SMAP_L2 products based on SMAP-derived loss

functions [Koster et al., 2017] and the production of lower-latency SMAP_L4 products using GEOS-5 forcing inputs without the benefit of gauge-based precipitation inputs.

Acknowledgments

Funding was provided by the NASA SMAP mission and NASA Terrestrial Hydrology Program via award 13-THP13-0022. Computational resources were provided by the NASA High-End Computing Program through the NASA Center for Climate Simulation at NASA/GSFC. The basin-scale datasets created and utilized in the analysis are available from the corresponding author.

6. Work Cited

Bretherton, C. S., M. Widmann, V.P. Dymnikov, J.M. Wallace, and I. Bladé (1999), The effective number of spatial degrees of freedom of a time-varying field, *J. Climate*, 12, 1990-2009.

Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, and M. H. Savoie (2012), EASE-Grid 2.0: Incremental but significant improvements for earth-gridded data sets, *ISPRS International Journal of Geo-Information*, *1*, 32-45, doi:10.3390/ijgi1010032.

Chen, F., W.T. Crow, P.J. Starks, and D.N. Moriasi (2011), Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, *Advances in Water Resources*, *34*(4), 526-536.

Crow, W.T., R. Bindlish, and T.J. Jackson (2005), The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall-runoff partitioning, *Geophysical Research Letters*, *32*, L18401, doi:10.1029/2005GL023543.

Crow, W.T., and D. Ryu (2009), A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals, *Hydrologic and Earth System Sciences*, *13*, 1-16.

Duan, Q., J. Schaake, V. Andréassian, S. Franks, G. Goteti, H.V. Gupta, Y.M. Gusev, F. Habets, A. Hall, L. Hay, T. Hogue, M. Huang, G. Leavesley, X. Liang, O.N. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener, and E.F. Wood (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and

309 third workshops (2006), *Journal of Hydrology*, *320*(1–2), 3-17,
310 doi:10.1016/j.jhydrol.2005.07.031.
311
312 Entekhabi, D., and Coauthors (2010), The Soil Moisture Active and Passive (SMAP) Mission,
313 *Proceedings of the IEEE*, *98*, 704-716, doi:10.1109/JPROC.2010.2043918.
314
315 Kerr, Y., and Coauthors (2010), The SMOS mission: New tool for monitoring key elements of
316 the global water cycle, *Proceedings of the IEEE*, *98*, 666–687,
317 doi:10.1109/JPROC.2010.2043032.
318
319 Kerr, Y.H., P. Waldteufel, P. Richaume, J. P. Wigneron, P. Ferrazzoli, A. Mahmoodi, A. Al
320 Bitar, F. Cabot, C. Gruhier, S. E. Juglea, D. Leroux, A. Mialon, and S. Delwart (2012), The
321 SMOS soil moisture retrieval algorithm, *IEEE Trans. Geosci. Remote Sens*., *50*(5), 1384–1403.
322
323 Koster, R. D., R. H. Reichle, and S. P. Mahanama (2017), A data-driven approach for daily real-
324 time estimates and forecasts of near-surface soil moisture, *Journal of Hydrometeorology*, in
325 press, doi:10.1175/JHM-D-16-0285.1.
326
327 Kumar, S.V., R.H. Reichle, R.D. Koster, W.T. Crow, and C.D. Peters-Lidard (2009), Role of
328 subsurface physics in the assimilation of surface soil moisture observations, *Journal of*
329 *Hydrometeorology*, *10*, 1534-1547, doi:10.1175/2009JHM1134.1.
330
331 Lei, F., W.T. Crow, H. Shen, R.M. Parinussa, and T.H. Holmes (2015), The impact of local
332 acquisition time on the accuracy of microwave surface soil moisture retrievals over the
333 contiguous United States, *Remote Sensing*., *7*(10), 13448-13465, doi:10.3390/rs71013448.
334
335 Lievens, H., G.J.M. De Lannoy, A. Al Bitar, M. Drusch, G. Dumedah, H.-J. Hendricks Franssen,
336 Y.H. Kerr, S.K. Tomer, B. Martens, O. Merlin, M. Pan, J.K. Roundy, H. Vereecken, J.P. Walker,
337 E.F. Wood, N.E.C. Verhoest, and V.R.N. Pauwels (2015), Assimilation of SMOS soil moisture
338 and brightness temperature products into a land surface model, *Remote Sensing of Environment*,
339 180, 292-304, doi.org:10.1016/j.rse.2015.10.033.
340
341 Liu, Y., W. Wang, Y. Hu, and W. Cui (2016), Improving the distributed hydrological model
342 performance in Upper Huai River Basin: Using streamflow observations to update the basin
343 states via the Ensemble Kalman Filter, *Advances in Meteorology*, 2016, 4921616,
344 doi:10.1155/2016/4921616.
345
346 Lin, Y., (2011), GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage
347 IV Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory.
348 http://data.eol.ucar.edu/dataset/21.093. Accessed January 2017.
349
350 Massari, C., L. Brocca, T. Moramarco, Y. Tramblay, and J.-F. Didon Lescot (2014), Potential of
351 soil moisture observations in flood modelling: estimating initial conditions and correcting
352 rainfall, *Advances in Water Resources*, 74, 44-53, doi:10.1016/j.advwatres.2014.08.004.
353

Massari, C., L. Brocca, L. Ciabatta, T. Moramarco, S. Gabellani, C. Albergel, and P. de Rosnay, S. Puca, and W. Wagner (2015a), The use of H-SAF soil moisture products for operational hydrology: Flood modelling over Italy, *Hydrology*, 2(1), 2-22, doi:10.3390/hydrology2010002.

Massari, C., L. Brocca, A. Tarpanelli, and T. Moramarco (2015b), Data assimilation of satellite soil moisture into rainfall-runoff modelling: A complex recipe?, *Remote Sensing*, 7(9), 11403-11433, doi:10.3390/rs70911403.

National Weather Servive (2007), NATIONAL WEATHER SERVICE INSTRUCTION 10 -1605. Published online at: https://verification.nws.noaa.gov/content/pm/pubs/directives/10-1605.pdf. Last accessed May 01, 2017.

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, and R. Bindlish (2016), *SMAP Enhanced L2 Radiometer Half-Orbit 9 km EASE-Grid Soil Moisture, Version 1*. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. Accessed January 2017. doi: http://dx.doi.org/10.5067/CE0K6JS5WQMM.

Parinussa, R.M., T.R.H. Holmes, N. Wanders, W.A. Dorigo, and R.A.M. de Jeu (2015), A preliminary study toward consistent soil moisture from AMSR2, *Journal of Hydrometeorology*, *16*, 932–947, doi:10.1175/JHM-D-13-0200.1.

Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, and J. S. Kimball (2016), *SMAP L4 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 2*. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. Accessed January 2017. doi: http://dx.doi.org/10.5067/YK70EPDHNF0L.

Silvestro, F., and N. Rebora (2014), Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain, *Journal of Hydrology*, *519A*, 1052-1067, doi:10.1016/j.jhydrol.2014.07.042.

U.S. Geological Survey (2016), National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), Accessed January 2017, at URL http://waterdata.usgs.gov/nwis/. doi: http://dx.doi.org/10.5066/F7P55KJN.

Vrije Universiteit Amsterdam (Richard de Jeu) and NASA GSFC (Manfred Owe) (2014), *AMSR2/GCOM-W1 surface soil moisture (LPRM) L3 1 day 10 km x 10 km descending V001*, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). Accessed March 2017. http://disc.gsfc.nasa.gov/datacollection/LPRM_AMSR2_DS_D_SOILM3_001.html.

Western, A. W., and R.B. Grayson (1998), The Tarrawarra data set: soil moisture patterns, soil characteristics and hydrological flux measurements, *Water Resour. Res*., *34*(10), 2765–2768.

Zhuo, L., and D. Han (2016), Could operational hydrological models be made compatible with satellite soil moisture observations?, *Hydrol. Process*., *30*, 1637–1648, doi:10.1002/hyp.10804.

Table 1. Attributes of study basins in Figure 1.

Basin Number	USGS Station No.	USGS Station Name	Basin Size (km^2)	Annual *P* (mm)	Runoff Ratio *Q/P*
1	07144780	Ninnescah River AB Cheney Re, KS	2,049	768	0.08
2	07144200	Arkansas River at Valley Center, KS	3,402	842	0.11
3	07152000	Chikaskia River near Blackwell, OK	4,891	896	0.19
4	07243500	Deep Fork near Beggs, OK	5,210	945	0.15
5	07147800	Walnut River at Winfield, KS	4,855	980	0.31
6	07177500	Bird Creek Near Sperry, OK	2,360	1025	0.23
7	06908000	Blackwater River at Blue Lick, MS	2,924	1140	0.29
8	07196500	Illinois River near Tahlequah, OK	2,492	1175	0.29
9	07019000	Meramec River near Eureka, MO	9,766	1187	0.28
10	07052500	James River at Galena, MO	2,568	1255	0.31
11	07186000	Spring River near Wace, MO	2,980	1258	0.27
12	07056000	Buffalo River near St. Joe, AR	2,148	1238	0.37
13	06933500	Gasconade River at Jerome, MO	7,356	1293	0.24
14	07067000	Current River at Van Buren, MO	4,351	1309	0.31
15	07068000	Current River at Doniphan, MO	5,323	1314	0.36
16	07290000	Big Black River NR Bovina, MS	7,227	1368	0.37

Table 2. The statistical significance of $\overline{R_s^2}$ differences sampled between all potential product pairs for a range of daily accumulation storm thresholds. Second row indicates $\overline{R_s^2}$ values taken from Figure 3. Arrows point to the product with the highest $\overline{R_s^2}$ for each pairing. Significance values are for a two-tailed hypothesis test.

5 mm/day

	AMSR2	SMOS_L2	SMAP_L2	NRv4	SMAP_L4	USGS SF
$\overline{R_s^2}$	**0.18**	**0.29**	**0.42**	**0.51**	**0.55**	**0.62**
AMSR2		↑ 96%	↑ >99%	↑ >99%	↑ >99%	↑ >99%
SMOS_L2			↑ >99%	↑ >99%	↑ >99%	↑ >99%
SMAP_L2				↑ 96%	↑ >99%	↑ >99%
NRv4					↑ 92%	↑ 99%
SMAP_L4						↑ 95%

15 mm/day

	AMSR2	SMOS_L2	SMAP_L2	NRv4	SMAP_L4	USGS SF
$\overline{R_s^2}$	**0.15**	**0.30**	**0.42**	**0.54**	**0.59**	**0.61**
AMSR2		↑ 98%	↑ >99%	↑ >99%	↑ >99%	↑ >99%
SMOS_L2			↑ 99%	↑ >99%	↑ >99%	↑ >99%
SMAP_L2				↑ 97%	↑ >99%	↑ >99%
NRv4					↑ 86%	↑ 90%
SMAP_L4						↑ 70%

25 mm/day

	AMSR2	SMOS_L2	SMAP_L2	NRv4	SMAP_L4	USGS SF
$\overline{R_s^2}$	**0.13**	**0.25**	**0.35**	**0.57**	**0.63**	**0.55**
AMSR2		↑ 96%	↑ >99%	↑ >99%	↑ >99%	↑ >99%
SMOS_L2			↑ 91%	↑ >99%	↑ >99%	↑ >99%
SMAP_L2				↑ >99%	↑ >99%	↑ 99%
NRv4					↑ 84%	← 61%
SMAP_L4						← 86%

35 mm/day

	AMSR2	SMOS_L2	SMAP_L2	NRv4	SMAP_L4	USGS SF
$\overline{R_s^2}$	**0.17**	**0.29**	**0.37**	**0.51**	**0.60**	**0.47**
AMSR2		↑ 78%	↑ 95%	↑ 99%	↑ >99%	↑ >99%
SMOS_L2			↑ 71%	↑ 97%	↑ >99%	↑ 95%
SMAP_L2				↑ 87%	↑ 98%	↑ 81%
NRv4					↑ 91%	← 65%
SMAP_L4						← 93%

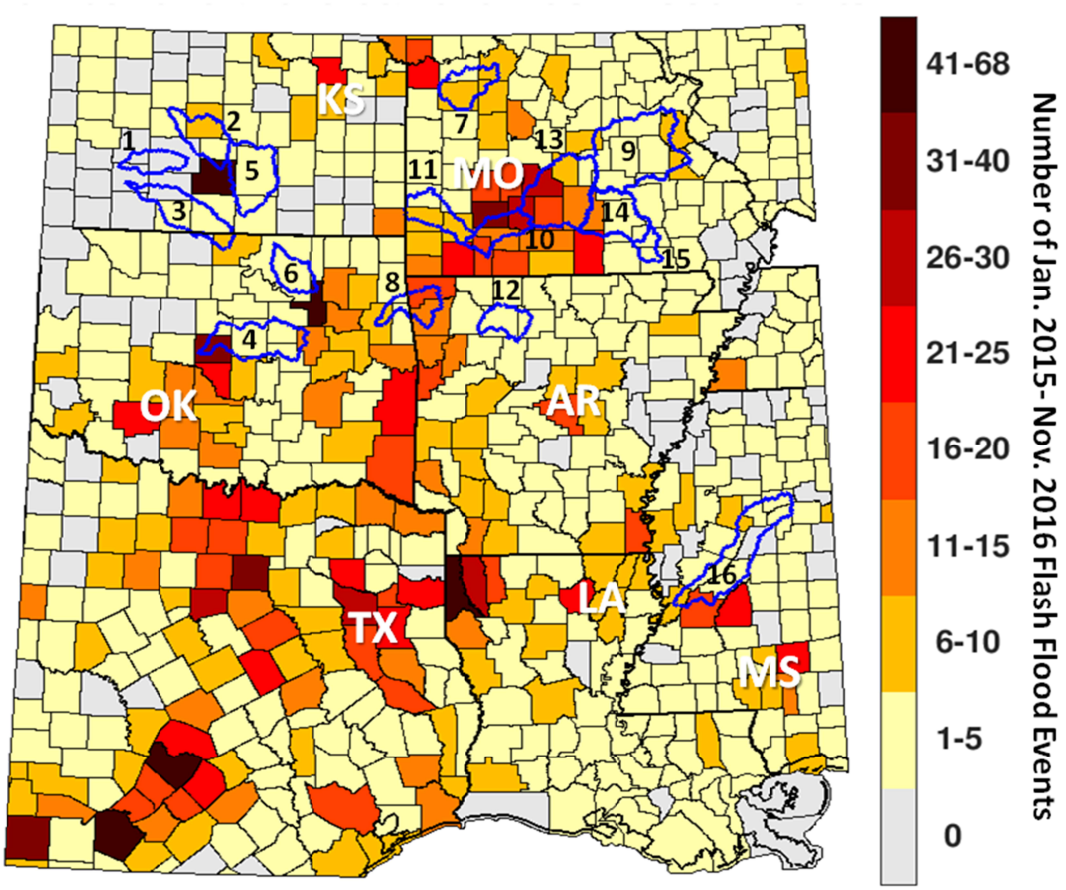


Figure 1. For a region of the south-central United States, boundaries (in blue) for our 16 medium-scale study basins overlain on a county-scale map of total number flash-flood events in the period Jan. 2015 to Nov. 2016. Identification of flash floods is based on the subjective reporting of major weather events by local weather observers to the United States National Weather Service (NWS) based on criteria described in NWS [2007]. Basins numbers correspond in the basin listing order given in Table 1, and individual US states are labeled.

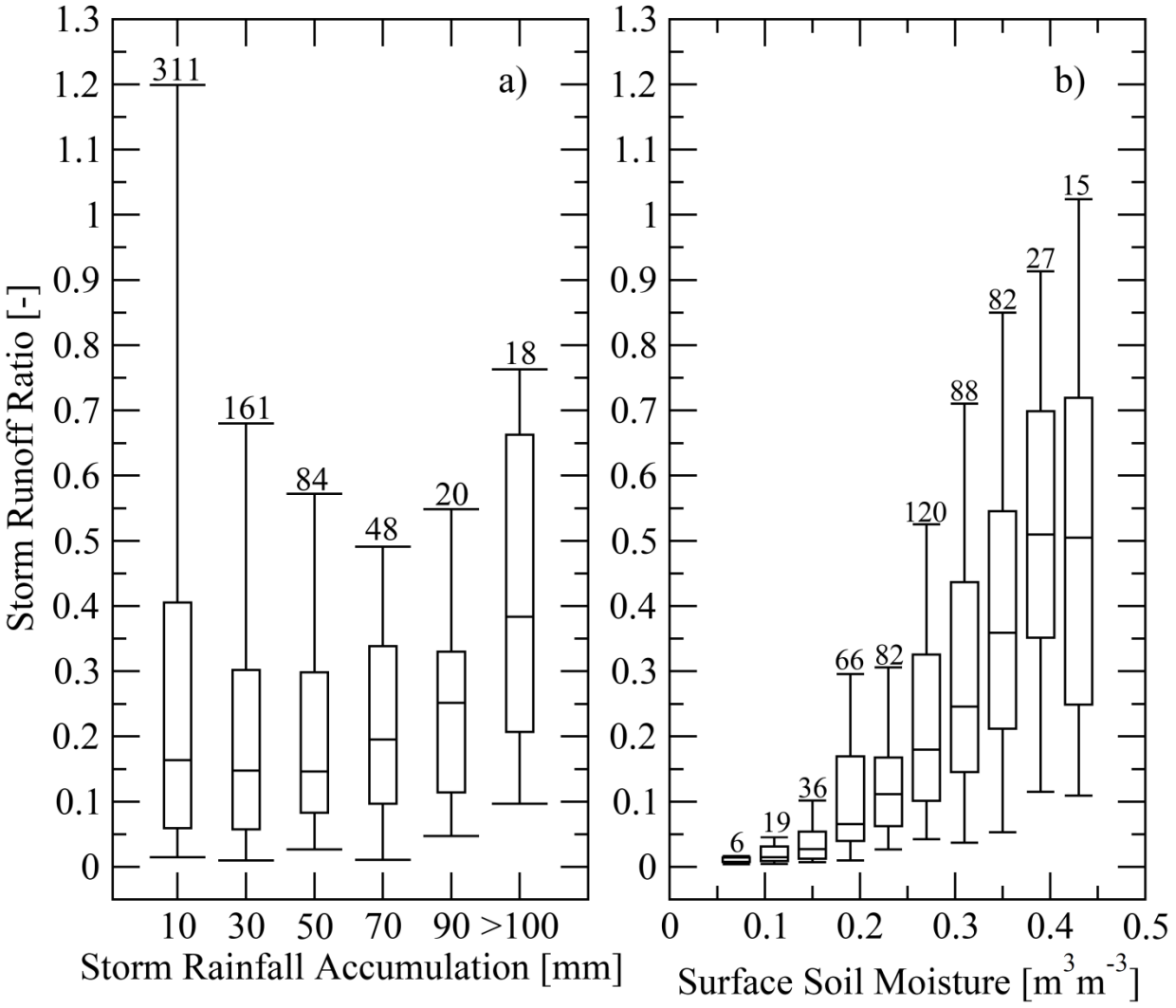


Figure 2. Box-plots (i.e., 5th, 25th, 50th, 75th and 95th percentiles) of storm-scale runoff ratio versus: a) total storm rainfall accumulation depths [mm] and b) pre-storm surface soil moisture [m^3m^{-3}] for storm events observed across all basins in Figure 1. In part b), pre-storm surface soil moisture is based on SMAP_L4 surface soil moisture estimates and events with accumulation depths less than 10 mm are excluded. Numbers represent total storm events described by each box-plot. Runoff ratios greater than one likely reflect measurement errors in estimates of storm total rainfall and/or stream flow used to determine the storm runoff ratio.

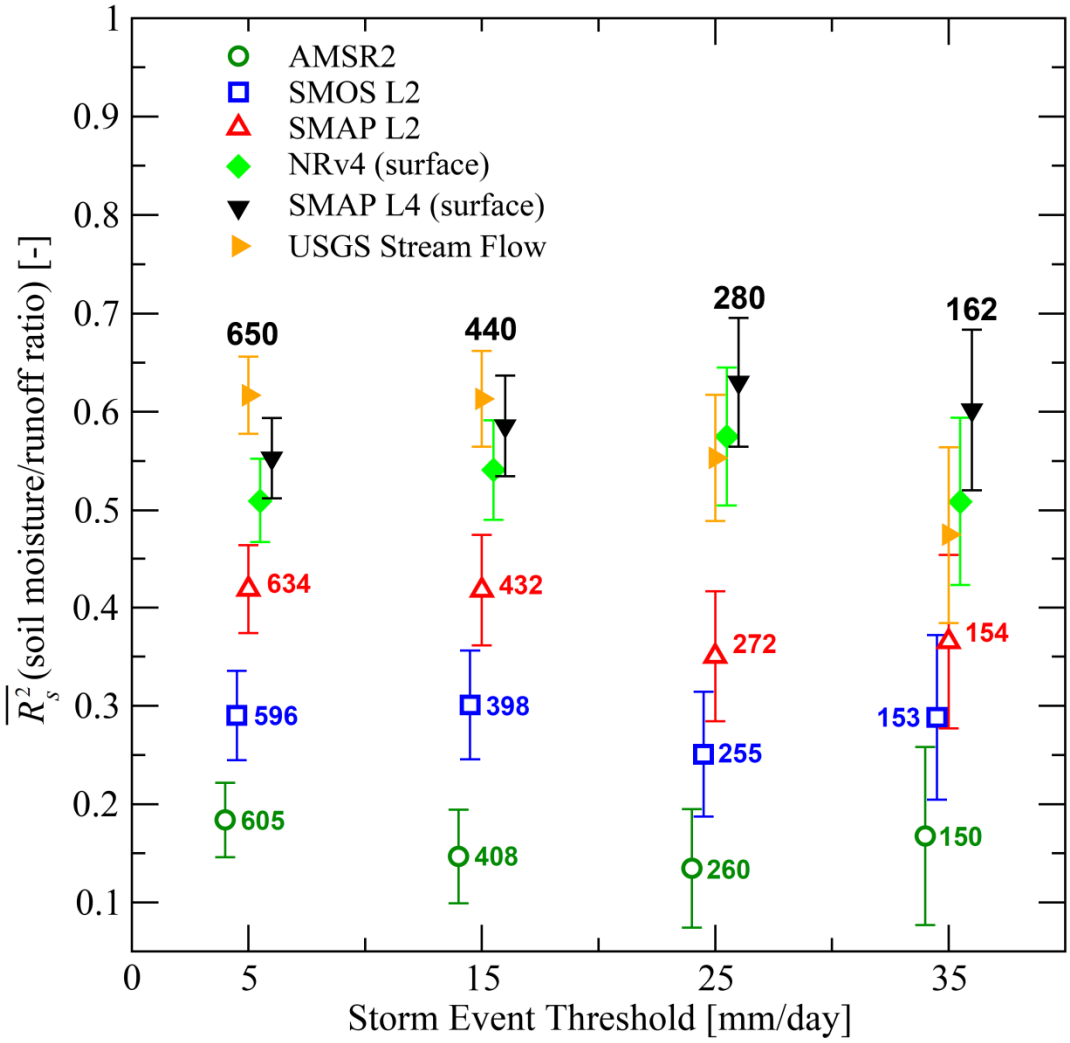


Figure 3. Spearman rank coefficient of variation $\overline{R_s^2}$ (between pre-storm soil moisture and storm-scale runoff ratio) versus storm event precipitation accumulation threshold for a range of soil moisture products (plus antecedent USGS stream flow). Error bars represent 95% sampling confidence. $\overline{R_s^2}$ is sampled in time within each basin and averaged across all 16 study basins (Figure 1). Numerical labels reflect the number of total storm events sampled to acquire $\overline{R_s^2}$. Symbols lacking individual numerical labels have complete temporal coverage and are based on the storm numbers indicated by the larger black numerals.